LA-UR-19-22947
Approved for public release; distribution is unlimited.

Title:	Ocean Current Segmentation at Different Depths and Correlation with Temperature in a MPAS-Ocean Simulation
Author(s):	Gospodnetic, Petra Banesh, Divya Wolfram, Phillip Justin Jr. Petersen, Mark Roger Hagen, Hans Ahrens, James Paul Rauhut, Markus
Intended for:	IEEE Scientific Visualization Conference (SciVis) Short Papers, 2018-10-21/2018-10-26 (Berlin, Germany)
Issued:	2019-04-02

Ocean Current Segmentation at Different Depths and Correlation with Temperature in a MPAS-Ocean Simulation

Petra Gospodnetic*
Fraunhofer ITWM

Divya Banesh†
University of California, Davis
Los Alamos National Laboratory

Phillip J. Wolfram
Los Alamos National Laboratory

Mark Petersen
Los Alamos National Laboratory

Hans Hagen
University of Kaiserslautern, Germany

James Ahrens
Los Alamos National Laboratory

Markus Rauhut
Fraunhofer ITWM

Abstract

When analyzing and interpreting results of an ocean simulation, the prevalent method in oceanography is to visualize the complete dataset. However, this can lead to data being missed or misinterpreted due to the distraction caused by the extraneous data of the simulation. Furthermore, when the data stretches over many layers in depth or over numerous time-steps, the ability to track attributes such as ocean currents becomes difficult due to the complexity of the data. We propose an image processing approach to simulation preprocessing for visualization purposes, which offers automation of ocean current tracking within a simulation and ocean current segmentation from the rest of the simulation data. Using the proposed approach, it is possible to automatically identify the most scientifically-relevant streams, extract them from the rest of the simulation and correlate their behavior with other simulation parameters.

Keywords: Ocean currents, ocean current segmentation, image processing

1 Introduction

A thorough understanding of ocean current behavior is vital for many aspects of ocean science. For instance, currents adjacent to a coastline (known as boundary currents) have a complex interaction with the thermodynamics of the ocean as well as the atmosphere near the land they are adjacent to [6]. These currents transport heat, salinity and biological life over long distances and can even influence the fishing trades of a region. In a recent paper, Lozier [8] highlighted the value of ocean currents and eddies for the meridional overturning circulation. When examining current behavior in a region and over time, it is important to identify the regions an ocean current occupies to assess how this influences surrounding regions. However, because currents are continuous and directed meandering corridors of water in the ocean, driven by wind, water density differences and tides, set within a turbulent ocean, such identification is not always easy.

We present an ocean current segmentation technique to separate regions corresponding to major currents from the remainder of the oceanic data. We demonstrate the value of our technique by applying the segmentation algorithm to data derived from an ocean simulation. We use the Model for Prediction Across Scales-Ocean (MPAS-Ocean) simulation, the oceanic component of the larger MPAS simulation [11] which determines interactions between land, ocean, atmosphere and sea-ice. The simulation is composed of an unstructured Voronoi mesh-grid, which can be varied to produce a higher resolution in certain areas and a lower resolution in others.

Our segmentation algorithm focuses on the kinetic energy values derived from this ocean simulation, where high values indicate

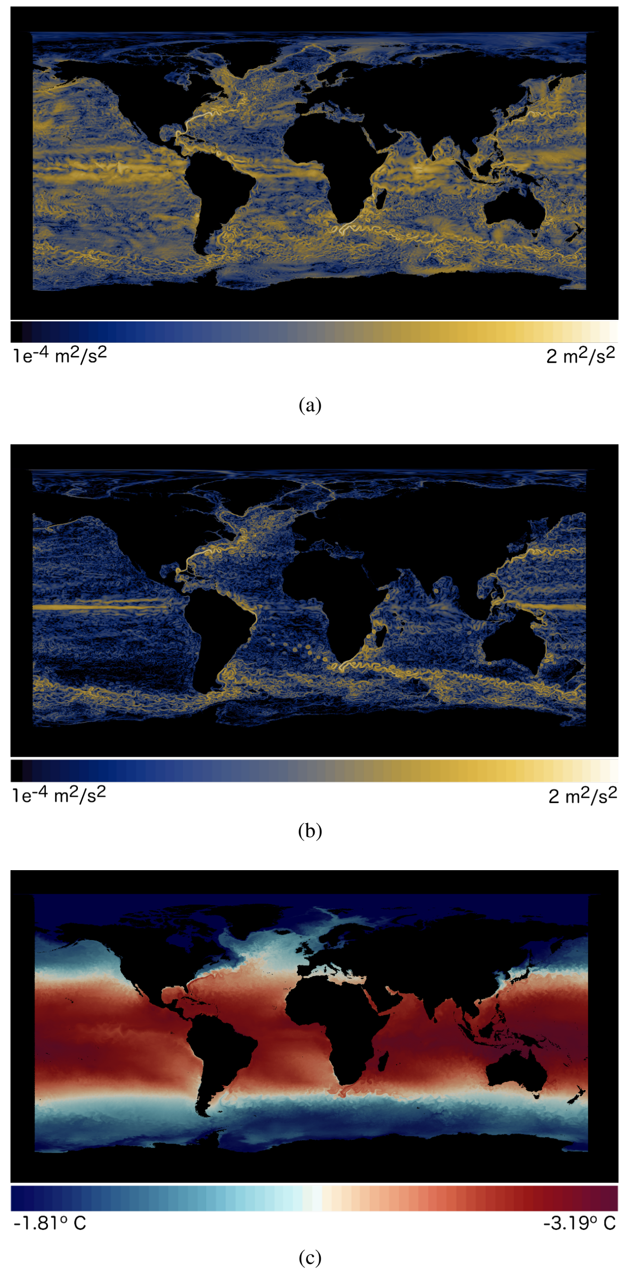


Figure 1: Color mapped MPAS-Ocean simulation visualization for different parameters; kinetic energy at surface (a), kinetic energy at 250 meters depth (b), temperature at surface (c).

*e-mail: petra.gospodnetic@itwm.fraunhofer.de
†e-mail: dbanesh@ucdavis.edu

presence of ocean currents. Currents have high translational energy and are most prominently detected in the kinetic energy parameter of the simulation. Ocean currents are present at all depths. To explore how data at different depths affect the results, we focus on the kinetic energy at the surface of the ocean, where currents are strongest, and at 250 meters, for comparison. The kinetic energy parameter is traditionally viewed on a logarithmic scale, between the range of $1e-2$ to $2m^2/s^2$, to maximize the values in the data. MPAS-Ocean is an ensemble simulation that provides various other parameters such as temperature, salinity, zonal and meridial velocity, etc. We analyze the data by projecting the global simulated data to a plane using the Mercator projection. Our data consists of 287 single channel images, each image five days apart in the MPAS-Ocean simulation. The resolution of each image is $2548x1310$ pixels and each pixel value is a projection of the data value of the simulation (within the above mentioned range) at a particular location in space. Our segmentation technique is applied to this raw data values, and the results are rendered at the end of the workflow to ensure accuracy.

As mentioned, it is to this set of projected data images that we apply our segmentation technique. *To extract and determine the network of currents in a region, we apply image processing techniques combined with principal component analysis (PCA) to find the corridors through which the strongest currents flow, excluding the noise in the dataset.* We can tune this *corridor identification algorithm* to include only the strongest currents or include a combination of currents at various strengths, depending on the user's requirements. We define *corridors* as areas encompassing the flow of an ocean current. Once we have identified these corridors, we map our results to temperature at sea surface, to find correlations, and showcase examples where the scientist can use an initial set of results for in-depth exploration. *This unique capability provided through our framework has been difficult to assess using traditional techniques [10], and has proven useful to the ocean scientist.*

Related work in image segmentation and region correlation with applications to climate sciences is expansive. Sukharev et al. [12] implemented clustering algorithms to find correlations and patterns in climate data sets. Results from the segmentation are then compared to other parameters of the climate simulation. While this is a more generalized approach, our algorithm is optimized for ocean current extraction. Chen et al. [2] presented techniques to identify correlations in time-varying data, but sample the dataset to reduce the time and effort required for correlation. Liebmann et al. [7] presented a hierarchical correlation clustering method for various 2D scalar fields. Our algorithm differs from these methods as we use image-based pixel information to derive our results.

In image processing, noise is considered to be a random variation of brightness. In our case the noise is the non-random data representing local turbulence, which we are not interested in when examining ocean currents. We are interested in strong ocean flows contributing to the global scale circulation and how this influences surrounding regions. By following the highest kinetic energy components in the simulation, it is possible to identify the aforementioned current flow corridors. The main aim of this work is to present a proof-of-concept on a new approach to ocean simulation parameter visualization, in order to explore its usefulness, and without optimization for in-situ analysis. It is certainly possible to apply it in-situ with future work. Furthermore, no error analysis is made since the presented tool is qualitative.

2 Current Segmentation Method

Since ocean currents are continuous directional movement of water, and they are passing through different regions of the ocean, we are proposing a segmentation method capable of ocean current corridor identification. The proposed current segmentation method is applied to a set of images, each a Mercator projection of the MPAS-Ocean simulation to a 2D plane. The dataset is a sequence of images, where each image corresponds to kinetic energy values at a given time-step, as can be seen from Fig. 2. Our simulation is comprised of 8km grid resolution at the equator, 18km resolution at the poles, and a gradual transition in-between. As a 15km resolution is often considered sufficient to identify most major features in an MPAS-Ocean simulation, these specifications satisfied our needs. All of the processing steps required for the segmentation are done for a single time frame, thus accessing the overall kinetic energy range relative to the highest and lowest values within the frame. When evaluating the results, we must also take into account that the values around the poles are distorted and further reduced due to the globe unwrapping process. We apply our technique to 2D images over the 3D simulation due to the ease of use a 402 MB database of images provides to the user over a 12 GB simulation. Ocean scientists most often focus on a particular region of the ocean for analysis because looking at an entire ocean simulation in detail is often impossible due to simulation size constraints. Image databases of simulations have proven to be effective for fast and efficient analysis [1]. By applying our technique to a set of images rather than a simulation, we expedited the analysis process while maintaining familiarity of the data to the scientist's typical workflow.

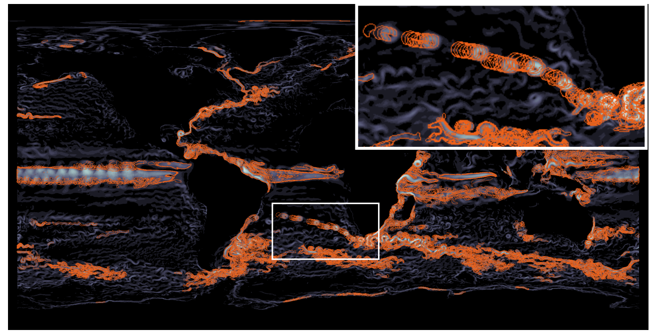


Figure 2: Kinetic energy in each time step is changing due to the movement of the currents. here, the movement is illustrated by computing and overlaying contours of the high kinetic energy areas for 15 sequential time-steps. The movement is the most apparent for eddies due to the consistent general direction of their trajectory.

2.1 Current Corridor Indicator Extraction

Kinetic energy (grey level) values are split into 4 classes using multi-level thresholding, where the thresholds are chosen to maximize the mutual information between the original pixel values and the pixel labels in the resulting image [9]. The new labels represent the strongest, medium and weak currents atop of the background. The most relevant information for the current flow tracking is stored within the strongest and medium current labels, and also big patches of weak currents. Therefore, we discard all of the small weak components. The remaining currents are used further as the foreground of a binary mask, on which is performed morphological (with a kernel size of $9x9$) closing followed by opening (kernel size of $3x3$) in order to connect smaller current components which are close by, as well as to remove the remaining currents which were left unconnected.

2.2 Current Component Orientation Estimation

Extraction of the biggest components results in a binary image containing corridor path indicators – a set of components which, when connected, would form significant current flows. The components are first labeled using the connected component labeling (CCL) algorithm [3], assigning every component a unique value and storing both its centroid and area.

After uniquely identifying the components, the next step is to connect them in a logical way. Decision on which components belong together and should be connected is a known and frequent problem in image processing. For an observer it may be obvious what belongs together, however when connections need to be described formally, the complexity of the problem arises. We decided to connect the components using orientation measure obtained from a principal component analysis (PCA) method [5].

PCA is a statistical method providing dimensionality reduction and is conducted in four steps: centering the data, calculating the covariance matrix, computing the eigen-decomposition and projecting the data. For the purpose of orientation measure, ontly the first three steps are utilized. Every labeled component is considered to be a set of 2D points, which are first centered by calculating and subtracting the mean for both axis. Further calculated covariance matrix, reveals the extent to which corresponding elements move in the same direction and it is geometrically interpreted using eigen-decomposition into eigenpairs $(\vec{v}, \lambda)$ i.e. eigenvectors $\vec{v}$ and their corresponding eigenvalues λ. Eigenpairs emphasize the direction of highest and lowest movement of the elements being analyzed, because of which we use the maximum eigenpair as a description of label orientation.

The decision to use the component orientation as the connection metric is based on the nature of the data. When looking at the overall kinetic simulation and discussing it with domain scientist, we defined two patterns: currents and eddies. Currents exhibit linear direction tendencies, have direction consistency and emphasize ridges with smaller streams branching from them. Eddies emphasize rotational components, have consistent origins and move in circular motion, thus forming the clear donut shape in the visualization and exhibiting no directionality as an extracted component. Motion of currents and eddies is correlated, as all the eddies originate from a current by breaking off from the stream and continuing to drift across the ocean in general direction influenced at origin. Considering such current behavior, connection of neighbouring component of similar orientation is a logical next step, with a search neighborhood for each label being an elliptical approximation defined by radii a and b as $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. The radii are assigned values of label width corresponding to the direction of eigenvectors, making sure the neighborhood search prefers components lying in the direction of the evaluated label. A label is considered to be in the neighborhood if a part of it lies within the elliptical search area.

Discovered neighborhood labels are further evaluated by comparing the difference in their orientation to the central label. If the orientation difference is less than 45°, the labels are connected by a straight line, making sure that the next CCL pass will identify them as a single connected component.

2.3 Corridor Formation

We use a *skeleton-based approach* to corridor estimation; the main flow is approximated by a thin line and then creates the corridor by expanding the line. The resulting image after connecting the components of similar orientation introduced inhomogeneity into path components (holes), which would represent a problem if we were to try and find the corridor skeleton directly from it. Therefore, a convex hull is calculated for each of the components first. Convex hull algorithm is constructing, developing, articulating, circumscribing or encompassing a given set of points in plane by a polygonal capsule called convex polygon [4], thus neutralizing the inhomogeneities and providing strong indicators of the area of interest. The hulls are again connected by performing morphological closing with a kernel of size $51x51$. Size of the kernel was determined heuristically, and its scale is to be expected since the operations are no longer based on the scope of a single current but on the greater parts of the flow.

Morphological merging of convex hulls results in several connected components along the current flows, which are irregular in shape and size but all equally important. To store the information about the path, while discarding the size irregularities, skeleton is obtained. Finally, the flow corridor is obtained by morphological dilation of the skeleton, with a kernel of size $51x51$, creating a uniformly wide corridor, indicating the most important currents.

Component connection step has a significant effect on the end result. The current method connects label centroids by a straight line, which may result in a corridor shift, when relatively big components

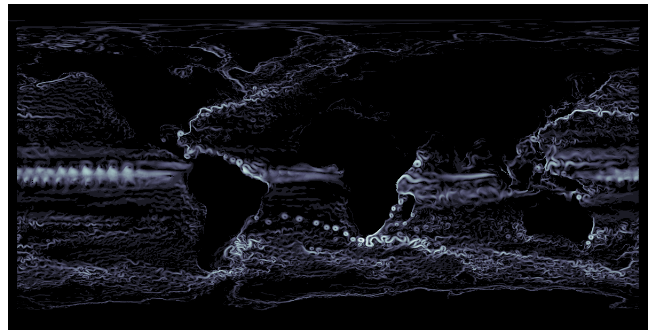


Figure 3: Cartesian projection of the kinetic energy data on the surface layer, enhanced by a logarithmic transform for improved visibility.

are compared because the centroids are not necessarily close to the component edges which should be connected ideally. In order to minimize the potential corridor shift effect and make sure to visualize the ocean current behavior around the most interesting components; the obtained corridor is merged with the binary mask covering the strongest and medium components

3 Results and Discussion

Using the approach elaborated in Section 2, we have successfully extracted the main ocean current flows out of the global kinetic energy data (Fig. 3), extended this to understand the behavior of the ocean temperature at sea surface height and kinetic energy at 250 meters ocean depth and visualized over time-steps in a video (for all videos, see additional material).

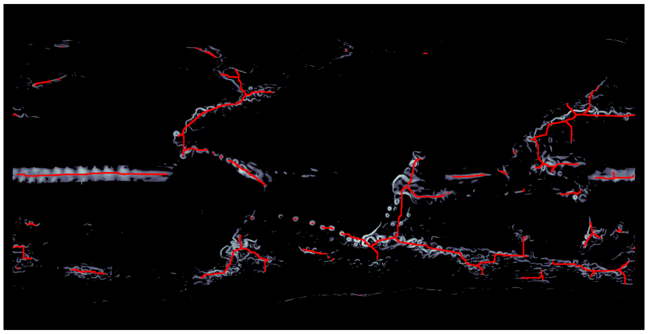
3.1 Current Segmentation

The corridor estimation approach was tested on two potential sets of corridor indicators, after multi-level thresholding results. First set contained only the strongest currents, whereas the second one contained biggest components obtained after the combination of strongest, medium and weakest currents.

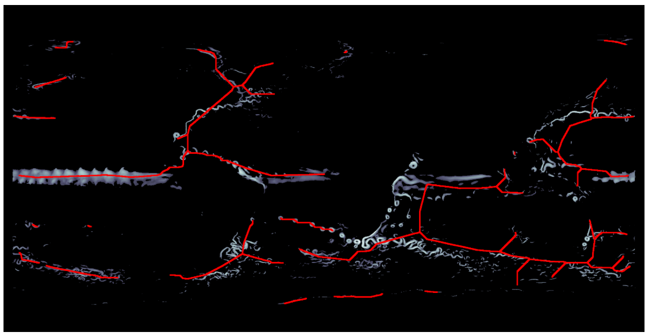
The proposed method assumes uniform spatial resolution over the entire dataset, which is not the case with this particular dataset. Furthermore, due to the Cartesian projection, the kinetic energies of the poles are distorted and therefore weakened, rendering the data invisible for extraction using our approach. Because of that, we can notice results fading out toward the poles. Such behavior is to be expected and is not considered a problem since the provided results serve to be a proof-of-concept and the method will be further used on the data with uniform resolution.

Using only the strongest components as corridor indicators, the method will fail to identify the Equatorial current, flowing between South America and Africa, as can be seen from Fig. 4(a), because it is relatively weaker than other currents, e.g. Agulhas Current (Cape of Good Hope). However, it is possible to overcome the problem by using the biggest components, as can be noted in Fig. 4(b). The corridor is computed for each frame and has successfully identified regions of interest within which we expect the ocean movement strong currents as well as the consequent weaker currents.

The results confirm the possibility of the corridor shift stated in Section 2.3, which can be noticed in Fig. 4(b), along the eastern coast of Africa (Agulhas Current). Success of the solution for the corridor shift impact minimization is also confirmed since the current is still visible after segmentation. With total run time of approximately 10 s/image, the proposed method in its current form is computationally expensive because the optimization was not the main focus during the implementation, but we focused to explore the results yielded by this approach. In order to reduce the possibility of a corridor shift and enhance the overall approach, we propose the following to be considered in the future work:



(a)



(b)

Figure 4: Application of the corridor computed using (a)only the strongest currents and (b) biggest components after combination of strongest, medium and weakest currents.

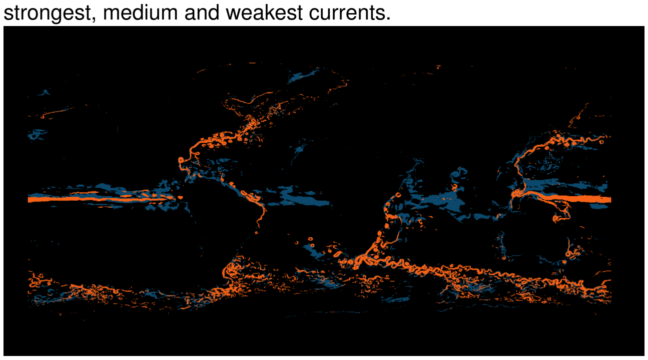


Figure 5: Overlay of the kinetic energy at 250 meters (orange) over kinetic energy on the surface layer (blue).

1) identification of the prominent edges in the label which should be connected, 2) evaluate and compare the local orientation of the discovered disconnected edges, 3) explore the possibility of using B-splines for corridor skeleton estimation and convex hull corners as their control points, 4) achieve temporal coherence by inter frame corridor skeleton interpolation.

3.2 Correlation of Currents to Kinetic Energy at Depth and Temperature

Using the proposed approach on the kinetic data, we have offered a possibility for spatial comparison of the surface currents to the currents at 250 meter depth, as shown in Fig. 5. The temperature in Fig. 6 has been extracted using the corridor obtained from the surface kinetic energy. Looking at the Fig. 1(c) it is possible to give global conclusions about the ocean temperature, but it is hard to say anything more about its exact correlation with the current flow. By extracting the current flow corridor and applying it onto the other simulation parameters such as temperature in Fig. 6 or kinetic energy at depth we offered a new possibility to observe other ocean current parameters, or compare the kinetic energy at different depths.

3.3 Domain Scientist Feedback

Our domain scientist collaborators were excited about the exploratory nature of this tool. One scientist noted that "Novel methods of analysis for ocean simulation data are extremely important. Computer science and image processing techniques are particularly relevant. I am interested in this application of segmentation to ocean modeling because velocity fields naturally parse into regions of energetic currents like the Gulf Stream, and slower return flows such as North Atlantic deep water." Another stated that "It is often hard to identify and understand key information in the entire global dataset without an approach or algorithm to focus exploration in local regions. For example, previous Lagrangian methods for identification of currents failed to provide any identification of the most important currents and furthermore failed to incorporate a multivariate analysis view of the flow, (e.g., they did not consider temperature or other scalar anomalies as in the present algorithm). The present method is unique and a great advancement because it provides an algorithmic and objective technique to identify the major ocean currents, which to my knowledge has not been previously accomplished to produced without use of individual 'expert knowledge'. The applications of this approach for future ocean currents in climate simulations is profound as expert knowledge (which may not exist for future climates) is not needed as the algorithm identifies major currents as opposed to heuristic identification by 'hand'."

4 Conclusion

The proposed segmentation approach using the biggest components of current intensities as corridor indicators confirmed the possibility to identify ocean current corridors enclosing the interesting currents as well as their context - surrounding weaker currents. The main advantage of such approach to current visualization is that, except for the morphological operation kernel size, we did not introduce any parameters which would have to be provided a priori. Further more, the domain scientist can focus only on the corridors as regions of interest when making an observation, without being distracted by weaker currents. The corridor extraction showed promising and useful results when applied to other simulation parameters such as temperature or depth. By using this frame-by-frame approach, it is possible to produce a video which provides the insight into the change of the observed parameters through time. The usefulness of 2D results points to the conclusion that, if the approach was to be conducted over depth, it can bring interesting new insights into the overturning circulation behavior for the domain scientists.

Acknowledgments

This research was supported in part by the Energy Exascale Earth System Model (E3SM) project funded by the U.S. Depart. of Energy Office of Science Biological and Environmental Research, Fraunhofer ITWM, and the University of Kaiserslautern.

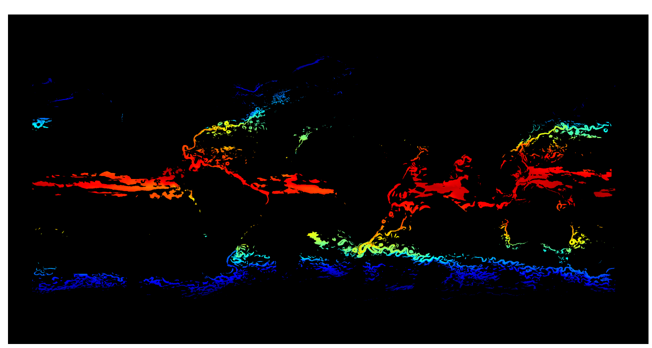


Figure 6: Visualization of the surface temperature parameter using corridors obtained from the surface kinetic energy.

References

[1] J. Ahrens, J. Patchett, S. Jourdain, D. H. Rogers, P. O'Leary, and M. Petersen. An image-based approach to extreme scale in situ visualization and analysis. *SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2014. doi: 10.1109/SC.2014.40

[2] C.-K. Chen, C. Wang, K.-L. Ma, and A. T. Wittenberg. Static correlation visualization for large time-varying volume data. In *Visualization Symposium (PacificVis), 2011 IEEE Pacific*, pp. 27–34. IEEE, 2011.

[3] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao. The connected-component labeling problem: A review of state-of-the-art algorithms. *Pattern Recognition*, 70, 2017. doi: 10.1016/j.patcog.2017.04.018

[4] M. A. Jayaram and H. Fleyeh. Convex hulls in image processing: A scoping review. *American Journal of Intelligent Systems*, 6, 2016. doi: 10.5923/j.ajis.20160602.03

[5] I. T. Jolliffe. *Principal Component Analysis*. Springer, New York, 2002. doi: 10.1007/B98835

[6] K. A. Kelly, R. J. Small, R. Samelson, B. Qiu, T. M. Joyce, Y.-O. Kwon, and M. F. Cronin. Western boundary currents and frontal air–sea interaction: Gulf stream and kuroshio extension. *Journal of Climate*, 23(21):5644–5667, 2010.

[7] T. Liebmann, G. H. Weber, and G. Scheuermann. Hierarchical correlation clustering in multiple 2d scalar fields. In *Computer Graphics Forum*, vol. 37, pp. 1–12. Wiley Online Library, 2018.

[8] M. S. Lozier. Deconstructing the conveyor belt. *Science*, 328(5985):1507–1511, 2010.

[9] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 1979. doi: 10.1109/TSMC.1979.4310076

[10] M. Petersen, X. Asay-Davis, A. Berres, Q. Chen, N. Feige, D. Jacobsen, P. Jones, M. Maltrud, T. Ringler, G. Streletz, A. Turner, L. Van Roekel, M. Veneziani, J. Wolfe, P. Wolfram, and J. Woodring. An evaluation of the ocean and sea ice climate of E3SM using MPAS and interannual CORE-II forcing, May 2018. doi: 10.5281/zenodo.1246339

[11] T. Ringler, M. R. Petersen, D. Jacobsen, M. E. Maltrud, and P. W. Jones. A multi-resolution approach to global ocean modeling. *Ocean Modelling*, 69, 2013. doi: 10.1016/j.ocemod.2013.04.010

[12] J. Sukharev, C. Wang, K.-L. Ma, and A. T. Wittenberg. Correlation study of time-varying multivariate climate data sets. In *Visualization Symposium, 2009. PacificVis' 09. IEEE Pacific*, pp. 161–168. IEEE, 2009.